

MPEG ビデオデータベースのための シーン格納機構のプロトタイプとその評価*

6Q-5

山下賢二, 進英二, 金子邦彦, 牧之内顕文

九州大学工学部情報工学科[†]

1 まえがき

近年, インターネット上で, ビデオを扱うことが可能となってきた. ネットワーク上で広く数多くの利用者がビデオを容易に扱うための技術として, ビデオデータベースが重要となっている. ビデオデータベースでは, ビデオの索引づけ, 分割格納, 検索などの研究が重要であり, その要素技術としてショット検出が必要である. 本論文ではMPEGビデオについて高速なショット検出法と, その処理速度・検出率の評価を報告する.

2 ビデオの格納と検索

ビデオを容易に取り出すには, 大量のビデオデータの中から必要なビデオを取り出すための機能(検索)が重要になる. 検索については, 題名や分野などのキーワードでの検索と, ビデオ内の実体やその動きによる検索が考えられる. いずれにしても, これらの検索を可能とするには, 1本のビデオをある単位(シーン)に分割し, シーン毎に実体や動きの情報を保持しなければならない. そこで我々は1本のビデオデータからシーンの切れ目(ショット)を検出し, ビデオをシーンの単位に分ける方法について研究を開始した. 従来, ショットの検出は画像処理の観点からさまざまな研究が行なわれ, 主に非圧縮の画像を対象としていた. 我々は, ビデオ格納・伝送のための標準的な符号化規格であるMPEGで圧縮されたビデオを復元することなくショット検出可能な手法を開発した.

3 MPEGについて

MPEGはピクチャ間の差異を抽出することで高い圧縮率を可能としている. 格納されたピクチャの符号化の種類には以下のような3つがある.

- Iピクチャ (イントラピクチャ)

他のピクチャの情報を使用せず, 独立して復号化

*A Shot Detection Algorithm for MPEG Video Database and its Performance Test

[†]Kenji Yamashita, Eiji Shin, Kunihiko Kaneko, Akifumi Maki-nouchi (Department of Computer Science and Communication Engineering, Kyushu University 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-81, JAPAN)

- Pピクチャ (予測符合化ピクチャ)

前のIピクチャまたはPピクチャを参照し, 復号化

- Bピクチャ (双方向予測符合化ピクチャ)

前と後のIピクチャまたはPピクチャを参照し, 双方の情報から復号化

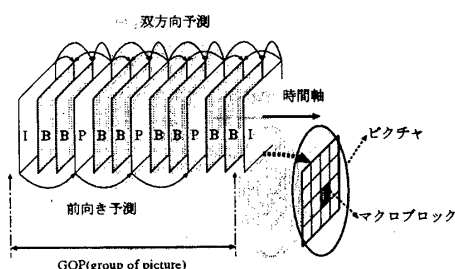


図1: MPEGの3つのピクチャ

また, MPEGの各ピクチャは格子状(16ピクセル×16ピクセル)のマクロブロックの集まりである.

このマクロブロックにはそれぞれマクロブロックタイプというパラメータが存在する. 符号化では, マクロブロックごとに以下の4つのタイプから圧縮率の最適なものが選択される.

1. イントラブロック (動き予測を用いないマクロブロック)
2. 前向き予測ブロック
3. 後向き予測ブロック
4. 双方向予測ブロック

我々はこれらのMPEGの特徴を利用して4章に述べるようなショット検出を行なった.

4 ショット検出

最初, 我々は, ピクチャのデータサイズが急激に増加してすると推定してショット検出を試みた. しかし, MPEGのIピクチャは常に動き予測を用いないのでデータサイズに大きな変化がなく, 検出不可能な場合があることが分かった.

そこで, 前後のピクチャの動き予測を使用するBピクチャに着目した. Bピクチャは, 前後双方向のピク

チャを参照して符号化されている。そのため、シーンの境界でないところ(図2)では、Bピクチャのマクロブロックのほとんどは、前後双方向の予測ブロックを用いている。そして、シーンの境界になると、その部分の前後のピクチャの相関関係がなくなり、前方向または後方向のみの予測を用いるブロックの使用率が增加する(図3)。そこで我々はBピクチャのマクロブロックタイプの割合を調べることで、ショット検出を行なった。

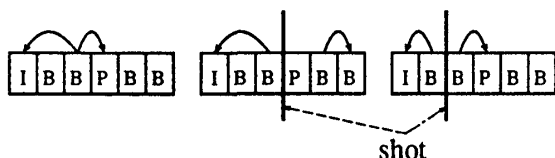


図2: ショットとショットでない場合のBピクチャ

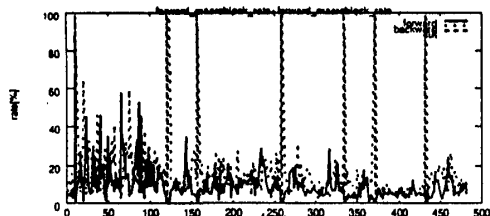


図3: Bピクチャの双方向予測使用率

ただし、この検出法はMPEGのフォーマットのうちBピクチャを使用しないものには適用できないという問題がある。

5 ショット検出の評価

4章に示した検出法の検出率と検出速度の結果を示す。

表1: 題材

タイトル	長さ	フレーム数	サイズ	フォーマット
ニュース	80秒	1434	6.7MB	IBBPBB...
ミニ英会話	59秒	1171	2.8MB	IBBPBB...
CG番組	27秒	484	1.8MB	IBBPBB...
プロ野球	560秒	10000	46MB	IBBPBB...

表2: 検出数

タイトル	ショット数	検出数	正答数	誤答数
ニュース	8	12	8	4
ミニ英会話	13	10	9	1
CG番組	7	7	7	0
プロ野球	179	228	157	71

表3: 検出率

タイトル	Recall	Precision
ニュース	100%	66.7%
ミニ英会話	69.2%	90.0%
CG番組	100%	100%
プロ野球	87.7%	68.9%

Recall= 正答数 / ショット数, Precision= 正答数 / 検出数

この結果より、Recall, Precisionともにほぼ70%以上であり、実用性がある。しかし、滑らかな変化は検出できない。逆に、画像の動きが激しい部分では、本来、ショットでない部分もショットとして検出してしまいう傾向がある。

表4: 検出速度

タイトル	長さ	フレーム数	検出時間	rate
ニュース	80秒	1434	123秒	1.53
ミニ英会話	59秒	1171	72秒	1.22
CG番組	27秒	484	36秒	1.33
プロ野球	560秒	10000	828秒	1.48

注)rateは(検出時間/ビデオの長さ)

従来のショット検出は、画像処理の観点から検出しており、精度は高いが、検出速度については未測定である。

この測定結果より、我々の検出法では実時間に近い速度で検出が可能であるといえる。

6 Webを利用したビデオ内容の表示

実際にショット画面を確認するために、ショット検出結果から、ショットの画面をホームページ化し、Netscape等のブラウザから一覧可能にするためのツールを作成した。これによって、ビデオの内容を容易に把握することも可能である¹。

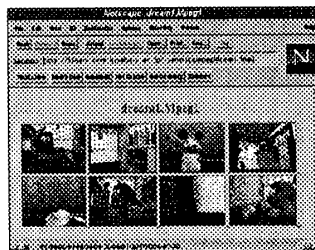


図4: Webを利用したショット画像の表示

さらに、CGI機能を利用して必要なシーンのみをクライアント側に転送するようにした。これはサーバ側で要求されたビデオの1部を切り出して転送している。

7 結論

実験の結果より、MPEGビデオを利用したショットの検出ではBピクチャのマクロブロックタイプに着目することが有効であることが分かった。また本手法は、検出速度については実時間に近い速度で検出可能であり、高速である。

参考文献

- [1] 岩成英一, 有木康雄, "DCT成分を用いた動画シーンのクラスタリング-カット検出の一手法-", 情報処理学会第47回全国大会

¹ショット検出とHTML作成プログラムは<http://hikaru.csce.kyushu-u.ac.jp/yamasita/mpegtool/index.html>で公開